2022 4th International Conference on Computer Communication and the Internet

Low Complexity Joint Power and Bandwidth Allocation for 3D Video SoftCast

1st Saqr Khalil Saeed Thabet
School of Information and Communication Engineering
University of Electronic Science and Technology of China
Chengdu, China
saqrthabet2004@yahoo.com

2nd Emmanuel Osei-Mensah
School of Information and Communication Engineering
University of Electronic Science and Technology of China
Chengdu, China
ecoseimensah@gmail.com

3rd Jamal N.A. Hassan
School of Information and Communication Engineering
University of Electronic Science and Technology of China
Chengdu, China
jamal1almelhany@gmail.com

4th Olusola Bamisile
College of Nuclear Technology and Automation Engineering
Chengdu University of Technology
Chengdu, China
boomfem@cdut.edu.cn

2022 4th International Conference on Computer Communication and the Internet (ICCCI) | 978-1-6654-6992-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/ICCCI55554.2022.9850266

***Abstract*—SoftCast transmission achieves a linear video quality transition commensurate with the wireless channel conditions, allowing it to avoid the cliff effect, unlike the digital video transmission system. When adopting SoftCast for three-dimensional video transmission, several issues arise. 1) Optimally allocate the power budget to texture and depth to minimize transmission distortion. 2) Select the appropriate number of texture and depth chunks to meet the bandwidth constraints. 3) Reduce the 3D video computation complexity. This paper aims to optimally solve the joint power and bandwidth allocation with low complexity. We derive a closed-form based on the power-distortion optimization problem to estimate the optimal power allocation ratio between the texture and depth map. Then to adapt to the bandwidth constraints. We first set the available bandwidth to be shared equally between texture and depth, this joint bandwidth allocation technique reduces the complexity, and to improve the overall 3D video quality, the joint power allocation is estimated at different bandwidth constraints. With low complexity, the proposed method achieves a graceful video quality transition with the improvement of channel conditions under bandwidth constraints and better performance than its counterpart default fixed-ratio power allocation between texture and depth.**

***Index Terms*—Power Allocation, 3D-DCT, Bandwidth Allocation, Distortion Minimization, Low Complexity**

I. Introduction

Three-dimensional video (3D video) has gained more attention in recent years as its application domains expanded to reach industry, education, healthcare, and entertainment. To represent the 3D content efficiently, Multiview video plus depth (MVD) [1] technique has been adopted. The growing capabilities of the current wireless communication networks (i.e., 5G and LTE) and the emergence of compact, high-quality 3D displays made it possible to deploy 3D video for home applications and handheld devices that can be used freely anywhere. Thus, 3D content-based applications on handheld devices are expected to gain a considerable market share in the coming few years. Nonetheless, some transmission challenges remain. First, transmitting a multi-view video over a wireless medium consumes substantial amounts of limited resources (power, bandwidth, and computation). Second, the unpredicted nature of the wireless transmission channels poses a higher requirement for 3D video transmission than traditional 2D video. Therefore, these issues must be addressed when designing any 3D video wireless transmission system. In 3D video wireless transmission, to deliver a fully immersive experience, considering Multiview video plus depth format, more than two neighboring views must be transmitted to the user. An array of closely spaced cameras captures these views of the same 3D scene from different angles. Based on the MVD format, every view is represented by texture and depth map frames. The transmitted views or the reference views are selected to generate a demanded virtual viewpoint by the user. Hence, the decoded texture video and depth map frames [2] are used at the receiver to synthesize that virtual viewpoint via depth image-based rendering (DIBR). The conventional digital transmission systems for 2D videos over wireless networks rely on Shannon's theorem in which the source encoder functions independently from the channel encoder [3] to compress the transmitted data and add error protection to cope with wireless channel conditions. However, the conventional digital transmission scheme does not suit the practical wireless transmission, where wireless channel conditions could vary drastically. When the channel quality falls beyond the minimum acceptable threshold, the reconstructed video quality degrades abruptly due to channel capacity mismatch and irreversible transmission errors, causing the collapse of the texture and depth decoding impacting the virtual viewpoint synthesis. Hence, the overall video quality of the reference and virtual views degrades abruptly. This phenomenon is known as the cliff effect [4]. Recently, SoftCast, an uncoded video transmission technique, has been proposed to overcome the conventional digital video transmission problems [5]–[7]. SoftCast design simply relies on ensuring linearity and avoiding quantization and entropy encoding. SoftCast transforms the transmitted signal into a series of real number coefficients corresponding to the pixel value using 3D-DCT transformation. These coefficients are then scaled linearly using a power allocation scheme to minimize the end-to-end distortion. Hence, the small wireless channel perturbation produces a small perturbation in the reconstruction quality of the received video. In this manner, SoftCast overcomes the cliff effect and smoothly scales the received video reconstruction quality within a wide range of wireless channel noise. Given the data division process implemented in SoftCast, neighboring coefficients with the same variance are gathered in blocks. The power allocation scheme implemented scales blocks based on their variance, which offers a different level of protection for the blocks. In the case of bandwidth constraint transmission, SoftCast can discard data blocks that include coefficients that do not contribute mainly to the video reconstruction quality. In the literature, many works suggested improvements to SoftCast. At the same time, both digital transmission and uncoded transmission can be incorporated into a Hybrid digital-analog framework to achieve efficient coding and robust adaptation. However, all schemes proposed apply to 2D videos, and a few works considered 3D videos. Yang et al. [8] proposed an uncoded wireless transmission that focuses more on depth maps. The scheme improves power allocation efficiency by considering view-synthesis distortion and scaling chunks formed by rearranging inter-block DCT coefficients. In [9], based on the 5D-DCT decorrelation, resource control algorithms were proposed to find the optimal solution for two problems: minimizing the distortion under a resource constraint and minimizing resource usage for a target distortion. Luo et al. [10] proposed a texture/depth joint power allocation method to achieve the optimal overall reconstruction quality for the reference and virtual views by solving the power-distortion optimization problem with respect to the view synthesis distortion. In [11], Saqr et al. proposed a joint power and bandwidth allocation method that estimates power allocation between texture and depth frames and then iteratively searches for the suitable number of texture/depth chunks for a given bandwidth constraint. In [12], Thabet et al. suggested incorporating a joint power allocation method to improve resource allocation problems to minimize the distortion under a resource constraint and reduce resource usage for a target distortion. Nonetheless, these works did not discuss how to allow the pseudo-analog 3D video transmission to accommodate the limited bandwidth allocated with low complexity. Different than 2D video transmission systems, which drop low variance chunks to match the bandwidth constraint without significantly degrading the quality, 3D video transmission system chunk discarding can not be applied in the same manner. In 3D videos, chunks that can be dropped from texture and depth frames must not drastically affect neither the reference views nor the virtual views. Additionally, 3D video systems that include multiple views computations and bandwidth allocation can add a burden to the system and increase time latency. Hence, this work suggests allocating the texture/depth power jointly to obtain the optimal overall 3D video quality under bandwidth and computation constraints. We propose a low complexity bandwidth allocation method to improve the pseudo-analog 3D video transmission performance under resource constraint conditions. Simulation results show a flexible performance over a wide range of wireless channels.

II. Joint Power and Bandwidth Allocation Method

It was shown in [10] that as the overall quality is measured after scaling 3D video texture/depth using different PAR, the overall quality differs in different PAR settings. Moreover, the video content decides the efficiency of the power allocation as the optimal PAR varies for different sequences. Hence, based on the power and quality relationship, a power-distortion model is utilized to find the optimal joint texture/depth PAR. Then a bandwidth allocation technique is proposed to adapt the transmission to the different bandwidth constraints.

A. Problem Formulation

Assume a 3D-DCT SoftCast transmission where M number of reference viewpoints are to be transmitted, the required total transmission Power P_{total} for texture and depth is indicated as:

$$P_{\mathit{total}} = \sum_{i=1}^{M}(P_{t,i} + P_{d,i}) \tag{1}$$

where $P_{(t,i)}$ and $P_{(d,i)}$ are the transmission powers for texture and depth of i^{th} reference viewpoint, respectively. Given L number of virtual viewpoints are to be rendered, then the total distortion of all the reference and virtual viewpoints D_{total} could be expressed as:

$$D_{\mathit{total}} = \sum_{i=1}^{M} D_{t,i} + \sum_{i=1}^{L} D_{v,i} \tag{2}$$

where $D_{(t,i)}$ and $D_{(v,i)}$ are the distortions of i^{th}reference view and the i^{th} virtual view, respectively. Assuming equal probability selection between display viewpoints, the transmission power for each reference view is the same. P_t is the same for all $P_{(t,i)}$ and P_d is also the same for all $P_{(d,i)}$. Thus, the power-distortion is simplified, and the joint power allocation problem to find the optimal texture/depth PAR that minimizes the total distortion of all viewpoints is:

$$\begin{gathered} D_{\mathit{total}}(P_t, P_d) = \arg \min_{(P_t, P_d)} \sum_{i=1}^{M} D_{t,i}(P_t) + \sum_{i=1}^{L} D_{v,i}(P_t, P_d) \\ \text{s.t. } M(P_t, P_d) < P_{\mathit{total}} \end{gathered} \tag{3}$$

B. Problem Solution

In order to optimally solve the joint power and bandwidth allocation with low complexity, a closed-form for the power-distortion relationship is formulated. Suppose an identical equal-size division is performed for both texture and depth. Hence, the total number of either texture or depth chunks is N (the total number of chunks in every reference view is

$2N$). Under a bandwidth constraint of T chunks, we assume splitting the available bandwidth equally between texture and depth frames to simplify the bandwidth allocation between texture and depth. The optimal scaling factor for sending $T/2$ texture chunks is:

$$g_{i,j}^{t} = (\lambda_{i,j}^{t})^{-1/4}\sqrt{\frac{P_t}{Z\sum_{j=1}^{T/2}\sqrt{\lambda_{i,j}^{t}}}} \tag{4}$$

And the optimal scaling factor for sending $T/2$ depth chunks is

$$g_{i,j}^{d} = (\lambda_{i,j}^{d})^{-1/4}\sqrt{\frac{P_d}{Z\sum_{j=1}^{T/2}\sqrt{\lambda_{i,j}^{d}}}} \tag{5}$$

where λ is the variance of each chunk, i and j refer to the j^{th} chunk of the i^{th} reference viewpoint and Z is the number of coefficients in each chunk. Then, the transmission distortions of the reference viewpoint for transmitting $T/2$ texture chunks and $T/2$ depth chunks are

$$D_{t,i} = \frac{Z^2\sigma_n^2}{P_t}\Big(\sum_{j=1}^{T/2}\sqrt{\lambda_{i,j}^{t}}\Big)^2 \tag{6}$$

$$D_{d,i} = \frac{Z^2\sigma_n^2}{P_d}\Big(\sum_{j=1}^{T/2}\sqrt{\lambda_{i,j}^{d}}\Big)^2 \tag{7}$$

According to the derivation in [10], the synthesis distortion of the i^{th} virtual viewpoints could be modeled as a linear combination of the approximated texture and depth distortions

$$D_{v,i} = \alpha_i\overline{D}_t + \beta_i\overline{D}_d + c_i \tag{8}$$

where α_i, β_i, and c_i are coefficients represent the model parameters. D_t and D_d are approximated from reference viewpoint transmission under different PAR settings and noise levels. By taking the distortion model (8), and the transmission distortions $D_{(t,i)}$ and $D_{(d,i)}$ into (3), the constrained optimization problem will take the following form

$$\begin{cases} D_{total}(P_t,P_d) = \arg\min\limits_{(P_t,P_d)} \dfrac{Z^2\sigma_n^2\sum_{m=1}^{M}\big(\sum_{j=1}^{T/2}\sqrt{\lambda_{m,j}^{t}}\big)^2}{M*P_t} \\ \quad \big(M+\sum\limits_{i=1}^{L}\alpha_i\big) + \dfrac{Z^2\sigma_n^2\sum_{m=1}^{M}\big(\sum_{j=1}^{T/2}\sqrt{\lambda_{m,j}^{d}}\big)^2}{M*P_d} \\ \quad \big(\sum\limits_{i=1}^{L}\beta_i\big) + \sum\limits_{i=1}^{L}c_i \\ \quad \text{s.t.}\,(P_t+P_d) < P_{total}/M \end{cases} \tag{9}$$

The convexity of the objective function is checked by taking the second-order derivative over P_t and P_d, as $\frac{\partial^2 D_{total}}{\partial P_t^2} > 0$, $P_t > 0$ and $\frac{\partial^2 D_{total}}{\partial P_d^2} > 0$, $P_d > 0$ and hence, the constrained optimization problem (9) can be solved using the Lagrangian Multiplier method,

$$\begin{aligned}\min_{(P_t,P_d)} J = & \frac{\big(M+\sum_{i=1}^{L}\alpha_i\big)Z^2\sigma_n^2\sum_{m=1}^{M}\big(\sum_{j=1}^{T/2}\sqrt{\lambda_{m,j}^{t}}\big)^2}{M*P_t} + \\ & \frac{\big(\sum_{i=1}^{L}\beta_i\big)Z^2\sigma_n^2\sum_{m=1}^{M}\big(\sum_{j=1}^{T/2}\sqrt{\lambda_{m,j}^{d}}\big)^2}{M*P_d} + \\ & \sum_{i=1}^{L}c_i + \mu\Big(P_t + P_d - P_{total}/M\Big)\end{aligned} \tag{10}$$

The optimal values P_t and P_d can be computed by solving the following set of equations

$$\begin{cases} \dfrac{\partial J}{\partial P_t} = \dfrac{\big(M+\sum_{i=1}^{L}\alpha_i\big)Z^2\sigma_n^2\sum_{m=1}^{M}\big(\sum_{j=1}^{T/2}\sqrt{\lambda_{m,j}^{t}}\big)^2}{M*P_t^2} + \mu = 0 \\ \dfrac{\partial J}{\partial P_d} = \dfrac{\big(\sum_{i=1}^{L}\beta_i\big)Z^2\sigma_n^2\sum_{m=1}^{M}\big(\sum_{j=1}^{T/2}\sqrt{\lambda_{m,j}^{d}}\big)^2}{M*P_d^2} + \mu = 0 \\ \dfrac{\partial J}{\partial \mu} = P_t + P_d - P_{total}/M = 0 \end{cases} \tag{11}$$

Then, the optimal PAR between texture and depth under T bandwidth constraint $PAR_T = P_t/P_d$ can be derived as

$$PAR_T = \frac{\sqrt{M+\sum_{i=1}^{L}\alpha_i}\sqrt{\sum_{m=1}^{M}\big(\sum_{j=1}^{T/2}\sqrt{\lambda_{m,j}^{t}}\big)^2}}{\sqrt{\sum_{i=1}^{L}\beta_i}\sqrt{\sum_{m=1}^{M}\big(\sum_{j=1}^{T/2}\sqrt{\lambda_{m,j}^{d}}\big)^2}} \tag{12}$$

C. Parameters Calculation

In order to estimate the optimal PAR_T value in (12), for specific bandwidth limitation, the parameters α_i, β_i, and c_i must be computed first for the L virtual viewpoints model. The model parameters are computed by pre-processing calculations. The reference viewpoint texture videos and depth maps are examined with different PAR settings, and noise levels (σ_j^2, PAR_j) to measure the reference viewpoint distortion and then perform linear fitting to obtain the corresponding model parameters. Since there are three parameters in the linear distortion model, at least three sets of data are required for the fitting solution. However, to improve the linear fitting accuracy to obtain model parameters, eight sets of (σ_j^2, PAR_j) were used. Four groups of different levels of noise, such as $(\sigma_{(t,1)}^2, \sigma_{(d,1)}^2)$, $(\sigma_{(t,2)}^2, \sigma_{(d,2)}^2)$, $(\sigma_{(t,3)}^2, \sigma_{(d,3)}^2)$, and $(\sigma_{(t,4)}^2, \sigma_{(d,4)}^2)$, where $\sigma_{(t,*)}^2$ is the noise variance applied to the reference viewpoint texture video, while $\sigma_{(d,*)}^2$ is the noise variance applied to the reference view depth map, in addition to two PAR settings. With these different sets, the average texture and depth distortion $\overline{D}_t^j, \overline{D}_d^j$ of the sets of reference viewpoints are obtained by computing the SSE between the original video sequence and the video sequence after applying the noise and PAR settings. For each virtual viewpoint i, $i \in 1, \cdots, L$, two video versions can be synthesized using 3D HEVC Test Model (HTM) v16.3 software [13]. The first is synthesized based on the original texture video and depth map. The second is synthesized based on the texture video and depth map after applying the different (σ_j^2, PAR_j) sets. Thus, with different (σ_j^2, PAR_j) sets, for each virtual view i, there are eight different sets of synthesis distortions $D_{v,i}^1, D_{v,i}^2, \cdots, D_{v,i}^8$. The linear regression method used to calculate the linear model parameters α_i, β_i, and c_i, considers solving the following system of equations:

$$\begin{cases} D_{v,i}^1 = \alpha_i \overline{D}_t^1 + \beta_i \overline{D}_d^1 + c_i \\ D_{v,i}^2 = \alpha_i \overline{D}_t^2 + \beta_i \overline{D}_d^2 + c_i \\ \quad\vdots \\ D_{v,i}^8 = \alpha_i \overline{D}_t^8 + \beta_i \overline{D}_d^8 + c_i \end{cases} \tag{13}$$

For the eight sets of (σ_j^2, PAR_j), we select four channels with SNR of 4 dB, 7 dB, 10 dB, and 13 dB, and two ratios of 1 and 6 for PAR. Considering one frame of each sequence to calculate the parameters reduces the pre-processing complexity. However, at the same time, as the large-size test sequences are divided into many chunks, we tend to consider eight frames of each sequence to calculate the model parameters. Algorithm 1 describes the whole procedure involved.

Algorithm 1 Joint Texture/Depth Power Algorithm

Input M, L, T
Output $PAR_T, \lambda_{m,j}^t, \lambda_{m,j}^d$

```
for m = 1 : M do
    for j = 1 : T/2 do
        Compute chunk variance λ^t_{m,j}, and λ^d_{m,j}
    end for
end for
for n = 1 : 8 do
    Apply different PAR settings and noise levels (σ_j^2, PAR_j) to texture and depth maps;
    Calculate the corresponding average distortions (D̄_t^n, D̄_d^n).
end for
for i = 1 : L do
    for n = 1 : 8 do
        Synthesize two virtual viewpoints using 3D-HEVC Test model (HTM);
        Calculate the corresponding synthesis distortion D^n_{v,i}.
    end for
    Compute α_i, β_i using the linear regression method.
end for
Calculate the optimal PAR_T
```

III. Simulation Results

Several simulations were conducted to assess the performance of the proposed low complexity joint power and bandwidth allocation for 3D video SoftCast.

Test Sequence: different reference 3D video sequences were used, which belong to two different resolution standards. The content of the standard video sequences Kendo, Balloons, Newspaper, PoznanStreet, and PoznanHall2 consist of captured real 3D scenes, whereas the standard video sequences Dancer, gtFly, and Shark are computer-generated 3D scenes with ground truth depth maps. The tested sequence configurations are listed in Table I.

Wireless Simulation Environment: Matlab simulation was implemented to conduct these experiments, and the OFDM physical layer was adopted with settings that match 802.11.a/g standard. The results are evaluated under AWGN channels.

TABLE I. 3D Video Test Sequences

Sequence	*Reference viewpoints*	*Virtual viewpoints*	*Resolution*	*Chunk width x height*
Balloons	1 and 5	2, 3, 4	1024x768	64x48
Kendo	1 and 5	2, 3, 4	1024x768	64x48
Newspaper	2 and 4	2.5, 3, 3.5	1024x768	64x48
Dancer	1 and 5	2, 3, 4	1920x1088	48x68
gtFly	1 and 5	2, 3, 4	1920x1088	48x68
PoznanHall2	5 and 7	5.5, 6, 6.5	1920x1088	48x68
PoznanStreet	3 and 5	3.5, 4, 4.5	1920x1088	48x68
Shark	1 and 5	2, 3, 4	1920x1088	48x68

A. Performance Evaluation of Joint Power Allocation

The performance of optimal PAR_T estimation with the joint power allocation method is assessed in terms of channel SNR = 4 dB, considering the average PSNR of all display views (reference and virtual views shown in Table I) as the overall 3D video quality. Table II lists the estimated PAR_T under different bandwidth constraints. In Fig. 1, the optimal PAR_T is estimated, assuming all chunks in the reference viewpoints are sent. The performance is compared to 14 PAR settings (PAR 1:1 to PAR 7:1 and extended to PAR 12:1 for the gtFly sequence). As Fig. 1 shows, each sequence has an estimated PAR_T that differs from the other sequences, close to the global optimal PAR_T, where the global optimum yields the best trade-off between the transmitted reference views quality and the rendered virtual views quality.

TABLE II: OPTIMAL PAR_T UNDER DIFFERENT BANDWIDTH CONSTRAINTS

Sequence	Estimated Optimal PAR_T			
	T=12.5%BW	*T=25%BW*	*T=50%BW*	*T=100%BW*
Balloons	2.80	2.74	2.61	2.23
Kendo	3.21	3.12	3.03	2.56
Newspaper	4.44	4.25	4.07	3.33
Dancer	3.66	3.64	3.70	3.65
gtFly	8.40	7.86	7.81	7.38
PoznanHall2	3.18	3.27	3.24	3.16
PoznanStreet	2.59	2.48	2.28	1.90
Shark	14.93	9.98	6.88	4.50

B. Performance Evaluation of Joint Power and Bandwidth Allocation

The performance of the proposed low complexity joint power allocation method under different bandwidth constraints is illustrated in Fig. 2. The channel range of the evaluation extends from 4 dB to 15 dB. The performance is evaluated under several bandwidth constraints (12.5%, 25%, 50%, and 100% of the total bandwidth). The proposed joint power allocation performance is compared to the same transmission system but with a fixed power allocation ratio between texture and depth (e.g.,1:1) under no bandwidth constraint. The proposed method's performance exceeds fixed PAR 1:1 over all the different test sequences, and it can reach a 1.8 dB gain for the gtFly sequence.

It can be seen that there is a performance gap between sending all the chunks in the reference view and sending a limited number of chunks according to the bandwidth constraints. At low SNR channels, the performance gap is within 1.6 dB on average between the minimum bandwidth constraint suggested and the total bandwidth. However, the difference in performance grows gradually as the overall quality requirement increases at high SNR channels, which transmission under low bandwidth can not match. Transmitting fewer chunks can result in significant errors and packet loss, degrading the overall quality.

C. Complexity Analysis

No iteration research was conducted in the proposed method to select the texture and depth chunks that match the bandwidth constraint requirement. The bandwidth allocation technique adopted in this method is to split the available bandwidth between texture and depth chunks equally.

As pointed out in [10] the complexity of the estimated optimal PAR under no bandwidth constraint was 96.25% lower than if the full search method was used to find the global optimal PAR. For the estimation method, given all reference viewpoints chunks to be transmitted, it pre-processes only eight times, including $8*2$ times distortion calculating the leftmost and rightmost reference views and $eight*L$ times virtual view synthesis. If bandwidth constraint is introduced, then after splitting the bandwidth between texture and depth equally, the method complexity will proportionate to the limited number of chunks included in the pre-processing and, at the same time, the joint power allocation estimation method will update the power allocation ratio PAR_T for every bandwidth constraint T.

IV. Conclusion

This work proposed a low complexity joint texture/depth power allocation under T bandwidth constraint for 3D video SoftCast. In order to estimate the optimal joint texture/depth power allocation ratio, a power-distortion optimization problem was designed. The problem is mathematically formulated and transformed into an unconstrained optimization problem then, by solving the problem using the Lagrangian multiplier, a closed-form of the optimal PAR is derived. Instead of proposing an iterative method to reach a suitable combination of texture and depth map chunks to meet the bandwidth resources constraint conditions, we simply suggested sharing the available bandwidth between the texture and depth map equally and estimating different PAR_T corresponding to the different bandwidth limitations to minimize the system distortion. The simulation results enumerate our method's effectiveness with a gain of 1.8 dB compared to the fixed PAR 1:1. Further, the proposed bandwidth allocation achieves a graceful overall 3D video quality over a wide range of wireless channel noise under bandwidth-constrained resources. In future work, the inter-view correlation will be considered to optimize the decorrelation process and further improve the

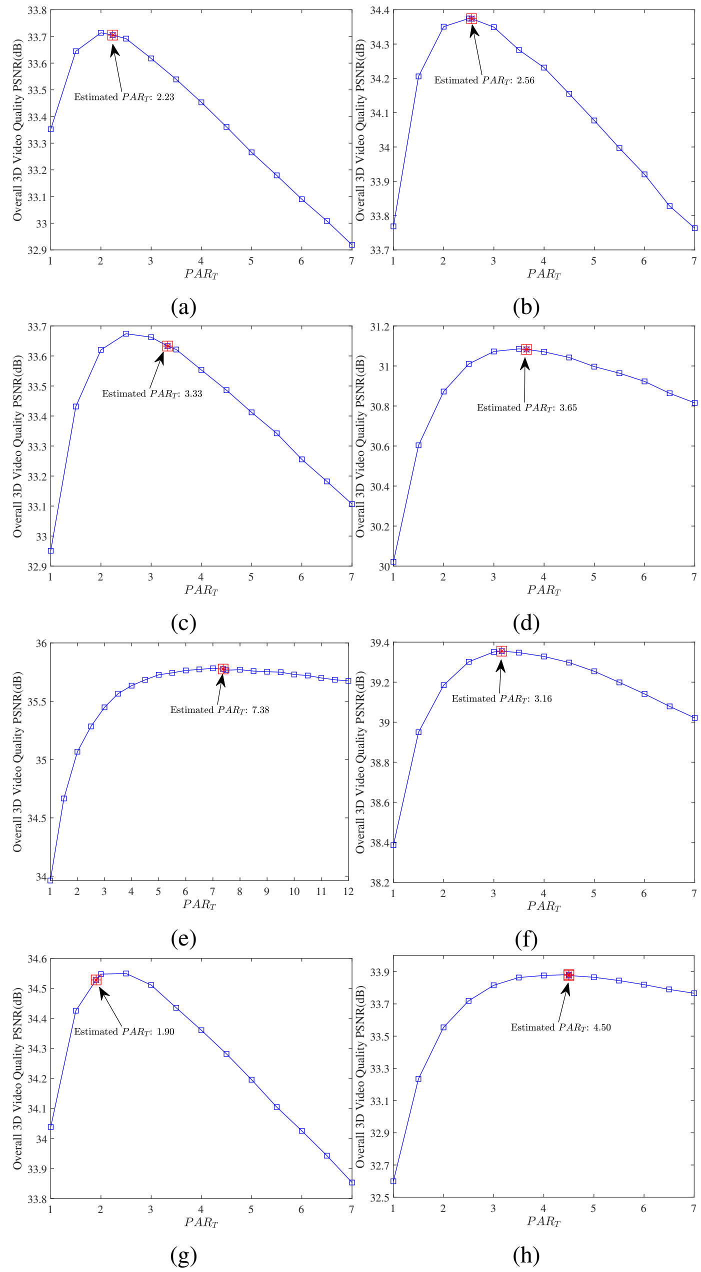
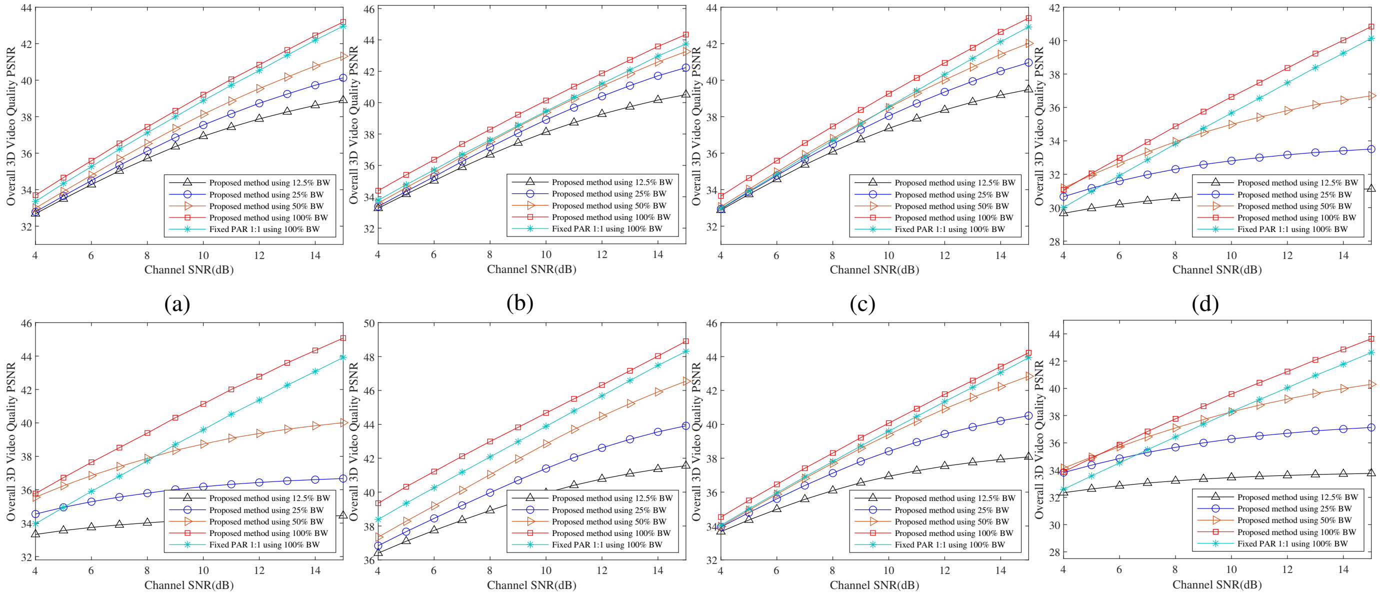


Fig. 1: Performance of different power allocation ratio PAR_T when the CSNR is 4dB. (a) Balloons; (b) Kendo; (c) Newspaper; (d) Dancer; (e) gtFly; (f) Poznanhall2; (g) Poznanstreet; (h) Shark.



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 2: Performance comparison between proposed power and bandwidth allocation and Fixed PAR 1:1. (a) Balloons; (b) Kendo; (c) Newspaper; (d) Dancer; (e) gtFly; (f) Poznanhall2; (g) Poznanstreet; (h) Shark.

3D video SoftCast transmission performance.

References

[1] P. Merkle, A. Smolic, K. Muller, and T. Wiegand, "Multi-view video plus depth representation and coding," in *2007 IEEE International Conference on Image Processing*, vol. 1. IEEE, 2007, pp. I–201.

[2] C. Zhu, Y. Zhao, L. Yu, and M. Tanimoto, "3D-TV system with depth-image-based rendering," *Architect Tech Challenges*, 2013.

[3] C. E. Shannon, "A mathematical theory of communication," *The Bell system technical journal*, vol. 27, no. 3, pp. 379–423, 1948.

[4] T. Kratochvíl and R. Štukavec, "DVB-T digital terrestrial television transmission over fading channels," *Radioengineering*, vol. 17, no. 3, pp. 96–102, 2008.

[5] S. Jakubczak, H. Rahul, and D. Katabi, "SoftCast: One video to serve all wireless receivers," *MIT Technical Report, MIT-CSAIL-TR-2009–005*, 2009.

[6] S. Jakubczak and D. Katabi, "SoftCast: One-size-fits-all wireless video," in *Proceedings of the ACM SIGCOMM 2010 conference*, 2010, pp. 449–450.

[7] ——, "A cross-layer design for scalable mobile video," in *Proceedings of the 17th annual international conference on Mobile computing and networking*, 2011, pp. 289–300.

[8] T. Yang, L. Luo, C. Zhu, and S. Tang, "Block DCT based optimization for wireless SoftCast of depth map," *IEEE Access*, vol. 7, pp. 29 484–29 494, 2019.

[9] T. Zhang and S. Mao, "Joint power and channel resource optimization in soft multi-view video delivery," *IEEE Access*, vol. 7, pp. 148 084–148 097, 2019.

[10] L. Luo, T. Yang, C. Zhu, Z. Jin, and S. Tang, "Joint texture/depth power allocation for 3-D video SoftCast," *IEEE Transactions on Multimedia*, vol. 21, no. 12, pp. 2973–2984, 2019.

[11] S. K. S. Thabet, E. Osei-Mensah, L. Lou, and C. Zhu, "Joint Power and Bandwidth Allocation for 3D Video SoftCast," in *6th International Conference on Digital Signal Processing*, Chengdu, China, 2022, in press.

[12] S. K. S. Thabet, E. Osei-Mensah, O. Ahmed, A. M. Seid, and O. Bamisile, "Resource optimization for 3d video softcast with joint texture/depth power allocation," *Applied Sciences*, vol. 12, no. 10, p. 5047, 2022.

[13] Fraunhofer Heinrich Hertz Institute, "3D High Efficiency Video Coding (3D-HEVC)—JCT-VC." [Online]. Available: https://hevc.hhi.fraunhofer.de/3dhevc